DOCUMENT RESUME

ED 193 317 TM 800 633

AUTHOR Jaeger, Richard M ; Wolf, Marian B.
TITLE An Experimental Comparison of Four Methods of Assessing Parents' Preferences for School System Goals.
INSTITUTION North Carolina Univ., Greensboro.
SPONS AGENCY National Study of School Evaluation, Arlington, Va.
PUB DATE Apr 80
NOTE 56p.; Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Community Attitudes; Educational Objectives; Elementary Education; *Intellectual Disciplines; *Parent Attitudes; *Questionnaires; School Surveys; *Test Construction; *Test Format; Test Items
IDENTIFIERS *Likert Scales; *Paired Comparisons

ABSTRACT

The traditional vehicle for collection of information from parents about a school's goals is a questionnaire that provides alternative goal statements. Two problems are frequently associated with such questionnaires: the goal lists suggest that a school can achieve all desired goals simultaneously; and many surveys use stimulus questions that are short, abstract, adn laden with educational jargon. Based on these general problems, an experiment was designed to compare responses of parents to four types of stimulus materials: three sets of materials in paired-choice format, varying in their level of specificity; and one set of materials in a Likert rating scale, at a broad level of specificity. Each of 224 respondents was asked to express his or her preferences for eleven curriculum areas. In the paired-choice format, reading was ranked highest; in the Likert format, reading and mathematics tied for the highest ranking. Music and art were ranked lowest. Rankings from the two formats were found to be substantially different. All paired-choice formats were more effective than the Likert format in allowing respondents to discriminate among their preferences. More concrete stimulus materials enhanced respondent discrimination among preferences, but did not result in substantial differences among preference orderings. (BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED193317

AN EXPERIMENTAL COMPARISON OF FOUR METHODS OF ASSESSING PARENTS' PREFERENCES FOR SCHOOL SYSTEM GOALS

by

Richard M. Jaeger

and

Marian B. Wolf

University of North Carolina at Greensboro

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Jaeger

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

© 1980, Richard M. Jaeger

This study was completed pursuant to a grant from the National Study for School Evaluation

TM 800633

TABLE OF CONTENTS

I.	Background and Problem	1
	A. The Ubiquity of School System's Surveys	1
	B. Typical forms of School Systems' Surveys	1
	C. Problems Investigated	4
II.	Method	7
	A. Overview	7
	B. Population and Sample of Respondents	9
	C. Instruments	17
	D. Interviewers	22
	E. Experimental Design	22
	F. Deviations from the Experiment as Planned	24
	G. Data Analysis Procedures	24
III.	Results and Discussion	26
	A. Relative Preferences for Eleven Curriculum Areas	26
	B. Comparability of Likert and Paired-Choice Results	29
	C. Discrimination of Curriculum Preferences as a Function of Stimulus Format	31
IV.	Implications for Practice in School Systems' Surveys	41
	A. Feasibility	42
	B. Choice of Stimulus Method	44
	C. Concreteness of Stimuli	45
Survey Instruments		Appendix A
Survey Documents		Appendix B

I. Background and Problem

A. The Ubiguity of School Systems' Surveys

In their attempts to be responsive to the demands of their communities, and in the face of diminished operating budgets, school systems often attempt to assess the relative values placed by parents and other taxpayers on various educational goals and objectives. The traditional vehicle for collection of information from parents and citizens is a questionnaire that provides alternative goal statements or listings of activities that could exist in schools. Respondents are asked to rate or rank these goal statements or activities, and thereby indicate the relative values they place on elements of the list (Savage & Jones, 1973; Ahmann, 1975; Eisner, 1972).

As in any scientific survey, a variety of methodological problems are associated with this traditional procedure for gathering information from parents and citizens. Foremost is the possibility that citizens' responses to questionnaire items might not reflect their true value positions on the activities and goals of the schools. This is a validity problem. Another potential problem is lack of consistency -- from one group of citizens to another, or from one time period to another. To the degree that they exist, these problems render the results of citizens' surveys less useful in planning responsive programs in schools.

To date, there have been no sound evaluations of the psychometric properties of school systems' citizens' surveys, nor have there been effectively designed assessments of the relative merits of alternative ways of seeking citizens' judgments. This study is an attempt to meet both needs.

B. Typical Forms of School Systems' Surveys

A review of the ERIC Documents File from 1966 to the present revealed a total of eighteen studies which used objective questionnaires to ascertain

parents' or citizens' preferences for educational goals or activities. Of these, fifteen used Likert rating scales (Beverly Hills, 1972; Blackwell, 1973; Bucks County, 1971; Crim et al, 1975; Eugene School District, 1975; Faunce et al, 1974; Horner & Martin, 1977; McGuire & Petrosko, 1978; McTeer et al, 1976; National Study, 1976; Preising, 1967; Rookey, 1976; Saskatoon Public Schools, 1973; Sawyers, 1976; Tacoma Public Schools, 1971), two used some sort of ranking technique (Arizona State University, 1969; Rankin et al, 1976), and one (Gilbert & Ellis, 1971) used a dichotomous response format. In addition to studies reproduced in the ERIC system, a Needs Assessment Compendium of Abstracts, compiled by the Rhode Island State Department of Education (1978), lists seventeen instruments with known response formats, designed to be used with parents or citizens in school system surveys. All seventeen used Likert rating scales. In addition, one instrument used checklists, two used ranking techniques, and one combined all three formats. It seems evident that Likert rating scales are by far the most widely used instruments for assessing citizens' and parents' preferences for educational goals or activities.

Various methods of collecting data were used in the studies reviewed, including telephone and person-to-person interviews, mail questionnaires, and questionnaires sent home with students. One study (Founce, 1974) used a method advocated by the Commission on Educational Planning of Phi Delta Kappa. In this method, respondents are gathered together, told to rate each of eighteen goals individually, and then assigned to a group of four so that they can reach concensus on ratings.

The studies reviewed included a broad range of specificity in the stimuli they employed. Among the studies which used rating scales, many stimulus

statements incorporated general labels for content areas, such as:

> "Our school is doing a good job in teaching children the basic skills. (Math, science, reading)." (National Study, 1976)

Other studies used more specific goal statements or activity descriptions, such as:

> "Our students analyze and evaluate reading selections (i.e. they distinguish fact from fiction)" (Rookey, 1967) and "Being able to add, subtract, multiply and divide numbers." (Preising, 1967)

A smaller number of the studies that employed rating scales provided both a broad goal statement and a more explicit description of activities:

> "HAS READING COMPREHENSION SKILL; DEVELOPS SKILL IN READING COMPREHENSION.
>
> Shows understanding of reading material by answering questions correctly, giving an oral or written report on reading assignment. Shows interest in improving by asking for extra help, reading more on his own, or showing a desire to improve." (Bucks County, 1971)

Both of the studies that involved ranking used broad educational goals as their stimuli. In one, (Rankin, 1976), respondents were given a list of eighteen goals and asked to list the three most important and the three least important. An example of these goals is "Improved reading skill."

The other survey which used ranking was conducted by Arizona State University (1969), and required respondents to choose the five most important goals from a list in which twenty goals were given. Sixteen additional blank spaces were provided so that the respondent could list others. An example of

the goals provided is "More effective teaching of language skills."

The study that offered only a dichotomous choice (Gilbert & Ellis, 1971) listed twelve content areas (English, Social Studies, etc.) and asked respondents to indicate whether or not they were satisfied (yes, no) with (1) course content, (2) quality of instruction, and (3) the standard of achievement for each area.

The examples given above illustrate that the statements to be rated or ranked are often extremely vague and may be interpreted quite differently by different respondents. In two of the studies reviewed, respondents even remarked that the items were ambiguous (Preising, 1967; Founce, 1974). Another study (McGuire & Petrosko, 1974), which used both general "supergoals" and more specific subgoals, reported that the ratings given to supergoals were often significantly different from the mean ratings given subsumed sets of subgoals. In fact, on a five-point scale, the difference between the average rating given a supergoal and the total mean rating given its subgoals was, in some cases, nearly a point. And in at least one case, it was over a point. Thus, although the developers of the instrument intended that each "supergoal" represent a given set of subgoals, this was not always the case.

C. Problems Investigated

As documented above, when school systems conduct surveys of parents or other citizens, they often seek opinions on the desirability of various educational goals and activities by using questionnaires that incorporate a Likert scale. Survey questions and response options are typically in one of two forms. One form poses questions of the type "Should the school system do X ?" and provides a four or five-option scale for response, with scale points ranging from "definitely no" to "definitely yes." An alternative form provides

a list of goals or activities, such as "The school system emphasizes development of functional reading skills," and asks respondents to choose options ranging from "strongly agree" to "strongly disagree."

This research was designed to investigate two problems associated with the content and format of stimulus questions used in school systems' surveys. First, questions and goal lists of the type illustrated above, together with their Likert response scales, suggest to respondents that a school system can achieve all desired goals simultaneously. Nothing stops the respondent from stating "definitely yes" or "strongly agree" in response to every stimulus question presented. But most school systems cannot afford to pursue all desired goals at the same time. Realities of budget and other resource limitations force tradeoffs among competing goals. Thus, the Likert form of survey questionnaire presents an unrealistic stimulus to survey respondents. Responses to such questionnaires are likely to be of little value in helping a school system's policy-makers choose among competing goals.

Second, most school systems' surveys use stimulus questions that are short, abstract, and laden with educational terminology that is likely to be unfamiliar to lay respondents. The goal statement illustrated above, "Development of functional reading skills," provides a good example. Abstractness and use of technical jargon can only be expected to reduce respondents' comprehension of the questions they are being asked to answer. The likely result is that respondents will interpret a given question in a variety of ways, contrary to the requirements of a high-quality survey (Moser and Kalton, 1972). The data derived from such surveys will probably reflect a mixture of respondents' true judgments or desires, and their misunderstanding or misinterpretation of the questions. In psychometric terms, the latter component constitutes

response error. The more error present in any survey, the lower will be survey reliability and the more difficult will be detection of true responses.

Based on these general problems, the research reported here examined four questions:

1. If respondents are asked to choose between elements of pairs of school system goal statements, will their preference ordering of such statements be similar to (or different from) the ordering resulting from use of a traditional Likert scale?
2. Will respondents discriminate between school system goal statements more effectively if they are asked to choose between elements of pairs of such statements, or if they state their preferences for goals in response to a traditional Likert scale?
3. Will use of more-concrete and explicit goal statements enhance respondents' abilities to discriminate among statements? (Or, equivalently, will use of more-concrete goal statements result in increased agreement among different respondents?)
4. How sensitive to the concreteness and explicitness of the goal statements they are asked to compare are respondents' relative preferences for particular statements?

As the foregoing discussion suggests, we embarked on this research with definite expectations about the answers to these questions. Our research hypotheses were as follows:

1. Presentation of goal statements in a paired-choice format (i.e. asking respondents to choose one or the other of two goal statements), rather than individually with an accompanying Likert scale, will result in a preference ordering that differs significantly from that produced by

the Likert scale format. In short, we hypothesized that a paired-choice format and Likert scale format present respondents with qualitatively different decision tasks.

2. Presentation of goal statements in a paired-choice format, rather than individually with an accompanying Likert scale, will result in significantly more-effective discrimination among respondents' preferences for various school system goals. For purposes of formulating school system policy on resource allocation, and for other policy-making purposes likely to be aided by knowledge of parents' and citizens' desires, we expected a paired-choice survey format to provide substantially more useful information.
3. Asking respondents to choose between more-explicit and concrete goal statements (in the paired-choice format) will result in greater discrimination among their stated preferences for various goals. Equivalently, more-concrete goal statements will result in greater inter-respondent agreement on a particular preference ordering of goals.
4. Variations among paired-choice goal statement formats (that is, using relatively more abstract or more concrete goal statements), will not result in substantial differences among respondents' preference orderings of goal statements.

II. Method

A. Overview

The study consisted of an experiment designed to compare responses to four types of stimulus materials presented in household interviews of parents of elementary school children. All of the stimulus materials sought parents'

judgments on the worth or value of eleven curriculum areas that constitute the elementary school curriculum of the Greensboro, North Carolina Public Schools.

Subjects (respondents) in the study were parents of children enrolled in kindergarten through grade six in the Greensboro Public Schools. Respondents were selected for participation in the study through a random sampling procedure, were invited to participate through a letter from the Greensboro Superintendent of Schools and the Dean of the School of Education at the University of North Carolina at Greensboro, and were interviewed in their homes at mutually agreed-upon times.

Three of the four types of stimulus materials were presented in a paired-choice format. That is, respondents were shown pairs of stimuli and were asked to select one or the other stimulus. These types of materials differed in the specificity with which they described goals or activities that represented the eleven curriculum areas in the Greensboro Schools' elementary curriculum. The fourth type of materials was presented in the form of a traditional five-point Likert rating scale. The stimuli (descriptions of curriculum activities) were identical to those used in the least specific of the paired-choice materials.

Equal numbers of respondents were randomly assigned to each of the four types of materials, and were interviewed independently. Depending on the type of materials used, the interviews required from fifteen minutes to just over an hour. Since it was likely that some respondents would not be literate, interviewers read all printed materials aloud at the same time respondents (may have) read them silently.

Methodological details are provided in the sequel.

B. Population and Sample of Respondents

1. Sampling Frame and Population

The sampling frame for the study was provided by the Greensboro Public Schools. It consisted of a computer tape containing the names and addresses of all students enrolled in the Greensboro Public Schools during the 1978-79 school year. Because the primary population of interest in this study was parents who had at least one child, who was not specially classified, enrolled in the Greensboro Public Schools in grades kindergarten through six during the 1978-79 school year, the available sampling frame contained a good bit of overregistration. It was necessary to eliminate those elements of the sampling frame that were outside the population of interest -- namely students enrolled in grades 7-12 and students who were classified as mentally handicapped. The constructed sampling frame (following the screening described above) and the population of interest thus matched very closely -- parents of students enrolled in grades kindergarten through six, not classified as mentally handicapped, during the 1978-79 school year. The only mismatch consisted of parents of students who had enrolled in the district since the sampling frame had been constructed, and parents of students who had left the school system since the sampling frame had been constructed. The latter group caused no operational problems since parents of students who had left the district could be randomly replaced, were they to be sampled.

2. Sampling Procedure

A sample of 224 respondents was desired for this study. Because the rate of cooperation among parents asked to participate in this study could not be anticipated in advance, it was decided to draw a random sample of 300. This was accomplished in two steps. First, the Statistical Package for the Social

Sciences (Nie, et al., 1972) was used to draw a sample of 900 students from the constructed sampling frame containing names of students enrolled in grades kindergarten through six, who were not classified as mentally handicapped. Names, addresses, and other identifying information for these 900 students were listed by computer. A linear systematic sample of 300 students was then selected from this list. Apart from minor problems of transiency, the final sample can be considered to be a simple random sample drawn from the population of interest. By design, names of parents and current telephone numbers should have been available for all sampled students. Unfortunately, this information was available for only a few of the students in the list. Since we planned to arrange interviews of parents in their homes at preassigned times, it was necessary to have telephone numbers for all respondents. Using students' last names and current addresses, we attempted to find current telephone numbers for their parents in the Greensboro Telephone Directory. This effort was successful for somewhat over 70 percent of the listed students. Telephone numbers for another ten to fifteen percent were located through the Greensboro City Directory. This directory indexes households by address rather than name. The Greensboro Public Schools provided current telephone numbers for another small percentage of the listed students. Those for whom current telephone numbers could not be found were not sampled. It is likely that this procedure resulted in a slight upward bias in the socio-economic level of our sample, and a slight racial bias as well. Representativeness of the final sample is explored in some detail below.

To assure representativeness, we would have liked to stratify the sampling frame on such variables as race and socio-economic level. Unfortunately, such data were not available in the sampling frame and stratification on variables likely to affect the final outcomes of the study was not possible.

3. Cooperation Rate and Replacement Procedure

A total of 327 letters inviting participation in the study were sent to sampled parents, and an additional nine parents were contacted by telephone without having received letters. Out of the 327 parents contacted by letter, there were 53 for whom no telephone number could be obtained, either through the telephone directory, the City Directory, or the Greensboro Public Schools. The project staff was unable to reach eight parents for whom telephone numbers were available, despite repeated attempts at various times of the day and evening. Twelve parents refused to be interviewed when they were telehponed, and an additional 39 could not be located at the times set for their interviews.

So, of 275 parents actually contacted, 224 agreed to participate in the study and were actually interviewed. The survey response rate was thus 224/275 or 81.5 percent. All parents who could not be contacted, refused to participate, or could not be located at the times set for their interviews were replaced by randomly sampling the constructed sampling frame.

4. Representativeness of the Sample

As seen in Table 1, the demographic characteristics of the respondents to this survey are generally similar to those of the population of families with children enrolled in the Greensboro Public School System, although some differences are apparent.

The relative frequency of white respondents to the survey is about eight points higher than the relative frequency of white students enrolled in the elementary grades in Greensboro. One possible reason for this difference is the large number of sampled parents for whom no phone numbers could be obtained. As was mentioned above, out of 327 parents contacted by letter, telephone numbers could not be obtained for 53, or 16.2%. It is reasonable to assume that a large proportion of parents for whom no telephone number could be found

Table 1. Comparison of Demographic Characteristics of Survey Respondents with those of the Population of Students in the Greensboro Public School System*

Variable		Percent of Survey Respondents		Percent of Population of Greensboro Public Schools
Race	% White	60.9		50.5 - 53.9
	% Nonwhite	39.1		46.1 - 49.5
Estimated SES	% Low	36.2		16.6 - 23.1**
	% Middle	53.8		49.7 - 54.2
	% High	10.0		26.4 - 30.2
Parents' Education		Mother	Father	
	% Finished High School or Less	46.9	42.3	58.0 - 59.2
	% Beyond High School	53.1	57.7	40.8 - 42.0

* Data for the population were available only by grade, for students enrolled in grades 1, 2, 3 and 6. Therefore, the figures in the table represent ranges of percentages for these grades. Data for the population were collected in Spring, 1978.

** SES data for the population were based on estimates of yearly family income. Low = less than $5,000, Middle = $5,000 to $15,000, High = over $15,000; these are not necessarily equivalent to "direct" SES estimates made for sampled respondents.

are of low socio-economic status (SES). And since previous research has shown a high correlation between race and SES, with nonwhites showing a greater tendency than whites to be of lower SES, it can further be assumed that the discrepancy in racial composition between our sample and the population can be attributed, in part, to our reliance on a telephone contact procedure.

To explore the potential effects of differences between the racial composition of the sample and the population on the research findings of this study, we computed Kendall's coefficient of concordance (a measure of subjects' abilities to distinguish among preferences for various curriculum areas or to agree with each other on a single preference ordering) for all white respondents and for all nonwhite respondents to one of the paired-choice formats. Weighted averages of the two coefficients of concordance were then computed, first weighting on the basis of the racial composition of the sample, and then weighting on the basis of the racial composition of the population. Comparison of the two weighted coefficients of concordance showed a difference of only .009, indicating that the disproportionate number of whites surveyed had little, if any, effect on the research findings of the study.

Apparent differences in the proportions of low, middle and high SES groups in the sample and population can be attributed largely to differences in measurement procedure. Figures for the respondents and for the population were based on subjective estimates. For the population, the SES categories were defined as estimated yearly family incomes of: (1) less than $5000, (2) $5000 to $15,000, and (3) over $15,000, respectively. In contrast, our interviewers were instructed to look around the homes of respondents and judge whether the family was of low, middle or high SES. Given the extremely

conservative income ranges used by the school system, it is understandable that the interviewers included more families in the low SES category and fewer in the high SES category than did the school system. It is likely that the interviewers included in the lower SES group, families who might earn more than $5000 per year, while limiting the upper SES group to families whose incomes might be estimated at well over $15,000 per year (the average annual income for a family of four now approaches $13,000).

The fact that the educational level of the survey respondents is somewhat higher than that of the population is consistent with observed differences in racial proportions, and might reflect the hesitance of those with less formal education to be interviewed. At the time our interview appointments were made, many parents had to be reassured by the survey staff that they would not be expected to bring to the interview any specific knowledge of school curriculum. In spite of this reassurance, many were unsure of their ability to respond intelligently. Although some of these people did eventually take part in the survey, a few refused, and others could not be reached at the appointed time.

Data in Table 2 allow comparison of respondents' demographic characteristics across interview formats. Although the statistics shown are not consistent with the categorical nature of the demographic data (except for the first variable), they do allow an assessment of the degree to which the four subsamples of respondents compared in terms of demographic characteristics. Using Tukey's Studentized Range procedure for multiple comparisons, only one significant difference was found among the seven variables examined. This was for the variable "Number of years the respondent has lived in Greensboro." For this variable, differences between means for three of the six possible

Table 2. Comparison of Demographic Statistics Across Samples of Respondents to four Interview Formats

Variable	Statistic	Interview Format: Paired-Choice Format: Activity-Type Statements	Paired-Choice Format: Activity Statements	Paired-Choice Format: Activity Statements with Photographs	Likert Scale Format: Activity-Type Statements	Tukey's Studentized Range Statistic
Number of children in Greensboro Public Schools	Mean	2.20	2.16	2.27	2.05	1.46
	Median	2.05	2.00	2.23	1.83	
	Standard Deviation	1.09	1.12	1.09	1.20	
Coded number of years respondent has lived in Greensboro * 1 = 1 year or less 2 = 2 to 5 years 3 = 6 to 10 years 4 = 11 years or more	Mean	3.13*	2.16	3.50	3.34	10.45**
	Median	3.22	2.00	3.76	3.72	9.20**
	Standard Deviation	.85	1.12	0.83	1.01	7.56**
Coded years of school completed by mother * 1 = less than 6 years 2 = 7 to 9 years 3 = 10 to 12 years 4 = 13 or more years	Mean	3.66*	3.39	3.43	3.45	3.45
	Median	3.78	3.39	3.50	3.46	
	Standard Deviation	0.55	0.59	0.63	0.57	
Coded years of school completed by father * 1 = less than 6 years 2 = 7 to 9 years 3 = 10 to 12 years 4 = 13 or more years	Mean	3.66*	3.40	3.56	3.41	3.11
	Median	3.78	3.44	3.69	3.53	
	Standard Deviation	0.55	0.63	0.60	0.71	
Coded race of respondents * 1 = White 2 = Black 3 = Oriental 4 = Other	Mean	1.43*	1.48	1.40	1.32	2.19
	Median	1.33	1.40	1.33	1.24	
	Standard Deviation	0.61	0.60	0.49	0.47	

Table 2. Comparison of Demographic Statistics Across Samples of Respondents to four Interview Formats

Variable	Statistic	Interview Format: Paired-Choice Format: Activity-Type Statements	Interview Format: Paired-Choice Format: Activity Statements	Interview Format: Paired-Choice Format: Activity Statements with Photographs	Interview Format: Likert Scale Format: Activity-Type Statements	Tukey's Studentized Range Statistic
Number of children in Greensboro Public Schools	Mean	2.20	2.16	2.27	2.05	1.46
	Median	2.05	2.00	2.23	1.83	
	Standard Deviation	1.09	1.12	1.09	1.20	
Coded number of years respondent has lived in Greensboro * 1 = 1 year or less 2 = 2 to 5 years 3 = 6 to 10 years 4 = 11 years or more	Mean	3.13*	2.16	3.50	3.34	10.45**
	Median	3.22	2.00	3.76	3.72	9.20**
	Standard Deviation	.85	1.12	0.83	1.01	7.56**
Coded years of school completed by mother * 1 = less than 6 years 2 = 7 to 9 years 3 = 10 to 12 years 4 = 13 or more years	Mean	3.66*	3.39	3.43	3.45	3.45
	Median	3.78	3.39	3.50	3.46	
	Standard Deviation	0.55	0.59	0.63	0.57	
Coded years of school completed by father * 1 = less than 6 years 2 = 7 to 9 years 3 = 10 to 12 years 4 = 13 or more years	Mean	3.66*	3.40	3.56	3.41	3.11
	Median	3.78	3.44	3.69	3.53	
	Standard Deviation	0.55	0.63	0.60	0.71	
Coded race of respondents * 1 = White 2 = Black 3 = Oriental 4 = Other	Mean	1.43*	1.48	1.40	1.32	2.19
	Median	1.33	1.40	1.33	1.24	
	Standard Deviation	0.61	0.60	0.49	0.47	

Table 2. (Continued)

Variable	Statistic	Interview Format: Paired-Choice Format: Activity-Type Statements	Interview Format: Paired-Choice Format: Activity Statements	Interview Format: Paired-Choice Format: Activity Statements with Photographs	Interview Format: Likert Scale Format: Activity-Type Statements	Tukey's Studentized Range Statistic
Coded estimates of respondents' SES * 1 = High social class 2 = Middle-class 3 = Lower or working class	Mean	2.28*	2.30	2.16	2.30	1.67
	Median	2.23	2.32	2.15	2.29	
	Standard Deviation	0.56	0.66	0.66	0.63	
Coded relationship of respondent to child * 1 = father 2 = mother 3 = male guardian 4 = female guardian	Mean	1.95*	1.96	1.91	1.96	0.93
	Median	1.96	1.97	1.95	1.96	
	Standard Deviation	0.41	0.38	0.29	0.50	

** Statistically significant, (α = .05)

pairs of subsamples (the three which included respondents who were interviewed using the Activity Statements paired-choice format) were found to be significant at the .05 level. Given the comparability of the subsamples on the other six demographic variables, it can safely be stated that any differences in research results cannot be attributed merely to differences between the subsamples of respondents.

C. Instruments

1. Types of Instruments

Four different types of stimulus materials were developed to ascertain parents' relative preferences for the eleven curriculum areas identified by the Greensboro Public Schools as those to be emphasized in grades kindergarten through six. Three of these stimulus types employed paired-choice formats and the fourth employed a Likert scale format.

Materials in the Likert scale format consisted of eleven general descriptions of types of activities designed to develop students' abilities in each curriculum area. Examples are:

> Activities designed to develop students' knowledge, skills and appreciation of music and dance through study or actual participation

> Activities designed to help students learn the language and concepts of basic science by reading and by observing and experimenting with the world around them; and to develop appreciation and respect for the world of nature.

Respondents were shown each stimulus statement individually, and were asked to indicate on a five-point scale, ranging from "Definitely Yes" to "Definitely Not," whether or not they wanted their child to take part in this type of activity at school. This interview format most closely resembled those typically used in the studies cited above.

Materials presented in the three paired-choice formats consisted of eleven stimulus statements presented to respondents in all fifty-five possible paired combinations. These stimuli differed among formats in terms of specificity and concreteness. The first paired-choice format used the general statements describing activity types that were used in the Likert scale format. The other two paired-choice formats consisted of longer, more explicit descriptions of specific activities in which students might engage, in order to develop their abilities in the eleven curriculum areas. Examples are:

AN EXAMPLE OF A MUSIC ACTIVITY

> Music is an important source of enjoyment and artistic expression. The schools try to develop students' abilities to appreciate music and to participate in creating music. In this activity, a professional musician comes to the classroom to perform, so that students can see what the musician's instrument looks like and how it is played. For example, a guitarist visits the classroom and plays for the students, so they can learn how the instrument is played and what kinds of sounds it can make. The guitarist teaches the students some songs, and they learn to sing along with the guitar.

AN EXAMPLE OF A SCIENCE ACTIVITY

> The world we live in contains solid objects (like rocks), liquids (like water and oil), and gasses (like air). Often, a substance can be changed from a liquid to a solid, or from a liquid to a gas, by heating it or cooling it. To learn about liquids, gasses and solids, and what happens to them when they are heated or cooled, children participate in supervised science experiments in their classes. They apply heat to blocks of ice, and observe that they turn to a liquid (water). They collect the water in a container, apply even more heat, and observe that the liquid changes into a gas (steam). They direct the steam into another container and cool it, observing that it changes back into a liquid (water). The teacher carefully explains the results of these experiments to the children, and explains to them how they are related to our weather and to causes of weather change, such as the sun's heat and the cooling of moisture in clouds to form rain.

The third paired-choice format used the same verbal material as the second, but, in addition, included photographs of students participating in the

activities described.

For the three paired-choice formats respondents were shown the stimulus items two at a time and were asked to indicate in which activity or type of activity (depending on the specific stimuli used) they would rather have their child participate. These paired-choice formats, unlike the more typically used Likert scale format, forced the respondents to make choices between alternatives, rather than allowing them to treat each stimulus independently. The increasing level of concreteness of stimuli used in the three paired-choice formats allowed comparison of preference orderings and consistency of response as a function of the specificity of stimuli.

A complete set of instruments is provided in Appendix A.

2. Development of Instruments

The first instruments developed for this study contained eleven general descriptions of types of instructional activities in elementary schools. These descriptions were used as stimuli in the Likert-scale interview format and in the first of the three paired-choice formats. They were based on program goals stated for grades kindergarten through six in the "Greensboro Public Schools Comprehensive Plan for Instruction," dated June, 1977. Since several program goals were listed by the school system for each curriculum area, an attempt was made to combine as many of these as possible into each of the activity type descriptions. For example, the "Healthful Living" goals stated as follows in the Comprehensive Plan:

> "Each student will:
>
> 1. Use health knowledge as it relates to the major structures and functions of the body for healthful living.
>
> 2. Develop an awareness of the relationship between health status and the major needs, sources of stress, and developmental characteristics of people throughout the human life cycle." (Greensboro Public Schools, 1977; p. 12)

became the activity type description:

> Activities designed to develop students' awareness of how to take care of their bodies to keep them healthy, how various body parts work, and the importance of maintaining a healthy body throughout life.

In the area of communication program goals, which included both reading and English-language arts, it was necessary to develop three activity type statements labeled "productive language," "receptive language" and "reading." The breadth of the school system's list precluded adequate coverage with fewer statements. In all other cases, the curriculum area identified in the Comprehensive Plan was translated into a single activity type statement. The eleven statements thus developed covered the curriculum areas of music, art, receptive language, productive language, reading, safety, physical education, health, citizenship, mathematics, and science.

When the eleven activity type statements were completed, they were submitted for review to the Greensboro School Leadership Team, composed of the Superintendent of the Greensboro Public Schools and his assistant superintendents. After gaining the approval of this group, these statements were used as the basis for developing more explicit activity statements, used in the two remaining paired-choice formats. To this purpose, the list of activity type statements was distributed to a group of teachers at a Greensboro elementary school, with the request that they suggest specific activities in each of the eleven curriculum areas which they might employ in their classrooms. The eleven specific activity descriptions were developed from the suggestions of these teachers.

In order to lend even more concreteness to the stimuli used in the third paired-choice format, the researchers went to a Greensboro elementary school and had students and teachers simulate the actual activities described, so that students could be photographed taking part in these activities. This was

accomplished through the cooperation of both the school administration and the students' parents, who, through letters sent home with their children and returned to the school, granted permission for their children to be photographed. These photographs were subsequently used in conjunction with the specific activity statements, to provide the most explicit set of stimuli employed in this study.

3. Demographic Information

Each of the four interview protocols provided for collection of demographic information, to be used to compare samples of respondents. Respondents were asked how many children they had enrolled in the school system, the length of time they had lived in Greensboro, the number of years of schooling they and their spouse had completed, and their relationship to the sampled child (this was determined at the beginning of the interview, but not recorded until the end). In addition, the interviewers were told to report, from their own observations, the race of the respondent and their subjective estimate of the social class of the family (high, middle or lower class). These subjective judgments were used, despite their obvious limitations, in order to avoid asking questions which the respondents might consider to be invasive.

For the purpose of comparing samples of respondents to the four types of interview formats, this method of data collection did not present a problem, since absolute validity of response was not of primary interest. Since random sampling of respondents was used, any bias should have been similar in magnitude and direction for all four interview formats.

Information on years of schooling, race, and social class was also used in examining the representativeness of the total sample in relation to the population. However, due to the subjective measurement of social class in both the population and the sample, as well as a difference in categories used to define

these classes, these data were of questionable utility for purposes of comparison.

D. Interviewers

Ten interviewers were used in this study. All were adults without previous experience in household interviews. Eight were women and two were men.

All interviewers participated in a two to three-hour training session prior to their initial interviews. They were informed of the general purposes of the survey (assessment of parents' judgments on the desirability of eleven curriculum areas in the Greensboro Public Schools elementary curriculum), but not of the research purposes of the study. Discussion of research objectives was avoided so as to minimize the effects of interviewer bias. All interview materials and procedures were reviewed in detail, questions on procedure were entertained, and interviewers were given the opportunity to practice interviewing procedures under the direction of the principal investigator and a research assistant. This practice followed a demonstration of appropriate procedures by the principal investigator and a research assistant. One pair of interviewers was videotaped as they practiced using survey materials, and their performance was later critiqued in the presence of the other interviewers.

As a part of interviewer training, potential problems were described by the principal investigator and a research assistant, and instructions for handling these problems were provided. Interviewers were instructed to telephone the survey coordinator following their initial interview, to discuss the experience, and report on any difficulties encountered. None of the interviews resulted in problems or difficulties that could not be resolved, and it appears that interviewer training was totally successful.

E. Experimental Design

The experiment conducted in this study was designed so as to minimize or eliminate several extraneous sources of variation. To eliminate possible

confounding of the effects of interviewers and interview stimulus materials, the experiment was designed so that each of eight interviewers would conduct seven interviews using each of four sets of stimulus materials. Thus, interviewers were used as blocks in a randomized block design. To eliminate potential order effects in the use of stimulus materials, two design features were used. First, interviewers were told to complete all interviews using one type of stimulus materials before moving on to interviews that used another type of materials. Second, two Latin square designs were used to determine the order in which stimulus materials would be used. By design, each type of materials would be used first by two interviewers, second by two interviewers, third by two interviewers, and fourth by two interviewers, thus eliminating the possibility of order effects. Interviewer numbers were assigned to interviewers randomly, thus minimizing the possibility of an interviewer by order-of-presentation effect. Random sampling of potential respondents and assignment of stimulus materials to respondents at random should have eliminated any possibility of systematic effects due solely to biased sampling of respondents.

In the paired-choice stimulus presentation formats, each stimulus was paired with every other stimulus, resulting in 55 pairs. The order in which stimulus pairs were presented to respondents was determined using a table of random permutations (Moses and Oakford, 1963). Seven different random orderings were used for each interviewer's seven interviews that employed a given type of stimulus materials. Identical orderings were used for the three types of materials presented in the paired-choice format so as to eliminate the possibility of an ordering-by-materials interaction.

For stimulus materials presented in the Likert scale format, stimuli were also arranged in random order using a table of random permutations. Again,

seven different random orderings were used for the seven interviews conducted by each interviewer.

F. Deviations From the Experiment as Planned

For the most part, the experiment was conducted just as described above. Because of illness or scheduling difficulties, three of the original interviewers could not complete all of the interviews assigned to them. One of the original interviewers completed 35, rather than 28 interviews as planned, and two additional interviewers were hired - one to complete four interviews, the other to complete three. This minor deviation from the planned experiment is unlikely to affect the ultimate results, although between-rater variance might have been increased by using ten interviewers rather than eight. Lack of complete balance in the assignment of types of stimulus materials to interviewers would affect the results only if there was some interaction between types of materials and interviewers. This possibility is highly unlikely.

G. Data Analysis Procedures

1. Conversion of Responses to Ranks

When presented with stimulus materials in one of the paired-choice formats, respondents indicated their relative preferences for curriculum areas by selecting one or the other element of each pair of stimuli. Basic response data thus consisted of a listing of these selections. A special FORTRAN computer program was written to count the number of times each respondent selected each of the eleven stimulus statements, when paired with each of the other ten stimulus statements. These eleven counts (ranging in value from zero through ten) were then ranked for each respondent, using the Statistical Analysis System (SAS Institute, 1979) computer package.

When presented with stimulus materials in the Likert format, respondents merely selected one of the scale values one through five for each of eleven

stimulus statements. The Statistical Analysis System computer package was then used to rank these scale values for each respondent. Determination of ranks was consistent with the assumption that respondents' scale-value selections constituted ordinal, but not interval, data.

2. Comparison of Preferences Across Stimulus Formats

To determine the consistency of respondents' relative preferences for various curriculum areas across stimulus formats, a two-step procedure was applied. First, the ranks assigned a given curriculum area were summed across all respondents presented with a given stimulus format. These rank sums were then ranked prior to completing the second step -- computation of Spearman rank correlation coefficients for each pair of stimulus formats.

3. Discrimination of Curriculum Preferences within Stimulus Formats

A variety of statistical procedures was used to determine respondents' abilities to discriminate among their relative preferences for the eleven curriculum areas. An alternative interpretation of the results of these procedures is that they indicate the consistency with which respondents presented with a given stimulus format agree on a particular relative preference ordering of the eleven curriculum areas.

To explore the null hypothesis that respondents preferred each of the eleven curriculum areas equally, Friedman's nonparametric two-way analysis of variance procedure was applied to the rank data computed for each of the four types of stimulus materials. On the assumption that these null hypotheses would be rejected for most (if not all) of the four types of stimulus materials, it was decided to search for statistically reliable differences in curriculum preferences by completing nonparametric analogs of Tukey's Studentized Range procedure for each pair of stimulus statements. This procedure allows

determination of preference differences for all pairs of stimulus statements, and by implication, determination of statistically reliable differences in ranking of curriculum area preferences.

A more-direct indicator of the consistency with which subjects agreed on a particular preference ordering of the eleven curriculum areas is given by Kendall's coefficient of concordance (W). This statistic is linearly related to Friedman's chi-square, the test statistic resulting from the two-way analysis of variance by ranks. Kendall's W can assume values in the range zero to one; zero indicates complete lack of agreement and one indicates complete agreement on a particular preference ordering. Kendall's coefficients of concordance were computed for each type of stimulus materials, so that the effect of stimulus materials on the consistency of respondents' preference orderings of curriculum areas could be examined.

III. Results and Discussion

A. Relative Preferences for Eleven Curriculum Areas

As described above, each of the 224 respondents in this study was asked to express his or her preferences for eleven curriculum areas. These curriculum areas: music, art, reading, productive language, receptive language, safety, physical education, health, citizenship, mathematics, and science, constitute the elementary school curriculum of the Greensboro Public Schools. Fifty-six respondents were presented with stimulus materials in each of four formats. Three of the formats were of the paired-choice type, and those respondents had to state which of two stimuli they preferred. The fourth format incorporated a Likert scale for response, and respondents had to rate each stimulus from one to five.

Since it was thought that the preferences of each respondent should be treated as ordinal data but not intervally-scaled data, their responses were converted to ranks. The totals of these ranks were then computed for the 56 respondents who were presented with each of the four stimulus formats. These rank totals were then ranked for each stimulus format, allowing comparison across formats. For each format the lowest rank total was assigned a rank of 1, the next lowest was assigned a rank of 2, and so on. Results are shown in Table 3.

Regardless of the stimulus format used, music was found to be the curriculum area least preferred by the respondents to this survey. Respondents presented with the paired-choice formats ranked reading higher than any other curriculum area. Reading and mathematics tied for the highest rank among respondents presented with the Likert format. Respondents presented with the totally verbal paired-choice formats and with the Likert format ranked art just above music. That is, it was second from the bottom in preference ordering. For those responding to activity statements with accompanying photographs in the paired-choice format, art ranked third from the bottom, just above safety.

As might be expected, mathematics was ranked very high by respondents to all stimulus formats. Those responding to activity-type statements in the paired-choice format, ranked it nine out of eleven. For those responding to activity statements, with or without accompanying photographs, in the paired-choice format, mathematics was ranked ten. For those responding to the Likert format, mathematics tied for top rank with reading. Thus at the extremes -- for the two lowest ranked curriculum areas and the two highest ranked curriculum areas -- there was a good bit of consistency across all stimulus formats.

Table 3. Ranks of Totals of Ranks Assigned Eleven Curriculum Areas by Four Samples of 56 Respondents. Based on Four Stimulus Question Formats.

Curriculum Area	Ranks of Rank Totals for 56 Respondents			
	Paired - Choice Formats			Likert Format
	Activity-Type Statements	Activity Statements	Activity Statements With Photos	Activity-Type Statements
1. Music	1*	1	1	1
2. Art	2	2	3	2
3. Reading	11	11	11	10.5
4. Productive Language	10	8	8	9
5. Receptive Language	8	9	9	3
6. Safety	4	4	2	8
7. Physical Education	3	3	4	5
8. Health	5	6	6	7
9. Citizenship	6	5	5	6
10. Mathematics	9	10	10	10.5
11. Science	7	7	7	4

*Note: To maintain consistency among data tables, the least preferred curriculum area was assigned a rank of 1, the next least preferred was assigned a rank of 2, etc. . .

For the paired-choice formats, the degree of consistency extended beyond the extremes. Respondents to this survey showed a definite preference for the basic skills curriculum areas, including reading and mathematics, and productive and receptive language. Science was ranked seven by respondents presented with each of the paired-choice formats, and was therefore preferred to some of the less-traditional curriculum areas such as health, physical education, safety, and citizenship. The rankings given curriculum areas not at the extremes were considerably different for those responding to the Likert format. Science ranked fourth from the bottom, and recpetive language was given the third lowest rank. The disparity in ranks between productive language and receptive language is unexpected, and is perhaps a methodological artifact. Latter analyses will show that preferences for many curriculum areas were barely distinguishable under the Likert format, and that differences between the rank totals of curriculum areas given a fairly high rank and those given a fairly low rank were actually quite small.

B. Comparability of Likert and Paired-Choice Results

Table 4 contains Spearman correlations between ranks assigned the eleven curriculum areas by those responding to the paired-choice formats and those responding to the Likert scale format. Although all three correlations are substantially greater than zero, and hypotheses of no relationship in the underlying population would be rejected at the .01 level of significance, all correlations are also considerably less than perfect. These results are consistent with previous observations based on data shown in Table 3. Two of the curriculum areas given the highest five ranks by those responding to activity-type statements in the paired-choice format are among those given the lowest five ranks by those responding to activity-type statements in the Likert format. The same observation is true for the curriculum areas given the highest five ranks by those responding to the Likert format. Two of those curriculum

Table 4. Spearman Rank-order Correlations Between Preference Orderings of Eleven Curriculum Areas, Resulting from Various Paired-Choice Stimulus Formats and a Likert-Scale Format*

	Likert-Scale Format
Paired-Choice Activity-Type Statements	.720
Paired-Choice Activity Statements	.688
Paired-Choice Activity Statements With Photographs	.606

* Ranks averaged across 56 respondents per format

areas are among those given the lowest five ranks by subjects responding to activity-type statements in the paired-choice format. Thus, a school system that adopted the policy of adding resources in the five curriculum areas most favored by parents or citizens, for example, would take very different policy actions depending on whether parents' or citizens' preferences had been determined through the paired-choice format or the Likert scale format.

Differences are even more striking when the stimulus materials, as well as the format, is changed. As might be expected, the correlation between ranks resulting from activity-type statements presented in the Likert scale format and the longer activity statements presented in the paired-choice format is somewhat lower than the corresponding correlation where the format, but not the wording, of the statements has been changed. When photographs are used with paired-choice activity statements, the correlation with the Likert scale results is smaller yet. These data tend to confirm our hypothesis that the paired-choice format and the Likert scale format would result in substantially different preference orderings of curriculum areas.

C. Discrimination of Curriculum Preferences as a Function of Stimulus Format

Our most strongly held hypothesis was that subjects responding in the paired-choice formats would do a substantially better job of communicating their relative preferences for various curriculum areas than would subjects responding in the Likert scale format. Data presented in Tables 5-8 bear on this research hypothesis, and we shall discuss these results in order.

Data in Table 5 apply to activity-type statements presented in the paired-choice format. They show the results of Friedman's test of the null hypothesis that members of the population represented by the 56 respondents sampled would

Table 5. Results of Friedman's Test and Nonparametric Analog of Tukey's Studentized Range Tests Applied to Average Ranks Assigned Eleven Curriculum Areas by 56 Respondents. Based on Presentation of Activity Type Statements in Paired-Choice Format.

Curriculum Area	1	2	3	4	5	6	7	8	9	10	11	Average of 56 Ranks
1. Music		N	-	-	-	N	N	-	-	-	-	2.643
2. Art			-	-	-	N	N	N	-	-	-	3.500
3. Reading				N	N	+	+	+	+	N	+	9.268
4. Productive Language					N	+	+	+	+	N	+	8.564
5. Receptive Language						+	+	+	N	N	N	7.304
6. Safety							N	N	N	-	-	4.348
7. Physical Education								N	-	-	-	3.875
8. Health									N	-	N	5.277
9. Citizenship										-	N	6.286
10. Mathematics											+	8.536
11. Science												6.402

Key:

- indicates curriculum area in row ranked significantly lower than curriculum area in column;

N indicates rankings of curriculum area in row and curriculum area in column were not significantly different;

+ indicates curriculum area in row ranked significantly higher than curriculum area in column;

Type I error rate (α) = 0.05.

Friedman's $\chi^2_r = 259.05$; $_{.95}\chi^2_{10} = 18.31$

Table 6. Results of Friedman's Test and Nonparametric Analog of Tukey's Studentized Range Tests Applied to Average Ranks Assigned Eleven Curriculum Areas by 56 Respondents. Based on Presentation of Activity Statements in Paired-Choice Format.

Curriculum Area	1	2	3	4	5	6	7	8	9	10	11	Average of 56 Ranks
1. Music		N	-	-	-	N	N	-	N	-	-	3.143
2. Art			-	-	-	N	N	-	N	-	-	3.509
3. Reading				+	N	+	+	+	+	N	+	9.813
4. Productive Language					N	+	+	N	+	N	N	7.491
5. Receptive Language						+	+	+	+	N	N	7.991
6. Safety							N	N	-	-	-	4.295
7. Physical Education								N	N	-	-	3.670
8. Health									N	-	N	5.643
9. Citizenship										-	N	5.036
10. Mathematics											+	8.768
11. Science												6.643

Key:

- indicates curriculum area in row ranked significantly lower than curriculum area in column;

N indicates rankings of curriculum area in row and curriculum area in column were not significantly different;

+ indicates curriculum area in row ranked significantly higher than curriculum area in column;

Type I error rate (α) = 0.05.

Friedman's $\chi^2_r = 267.48$; $_{.95}\chi^2_{10} = 18.31$

Table 7. Results of Friedman's Test and Nonparametric Analog of Tukey's Studentized Range Tests Applied to Average Ranks Assigned Eleven Curriculum Areas by 56 Respondents. Based on Presentation of Activity Statements with Photographs in Paired-Choice Format.

Curriculum Area	1	2	3	4	5	6	7	8	9	10	11	Average of 56 Ranks
1. Music		N	-	-	-	N	N	N	N	-	-	3.339
2. Art			-	-	-	N	N	N	N	-	-	3.741
3. Reading				+	N	+	+	+	+	N	+	9.875
4. Productive Language					N	+	+	N	N	-	N	7.036
5. Receptive Language						+	+	+	+	N	N	7.902
6. Safety							N	N	N	-	-	3.723
7. Physical Education								N	N	-	-	3.786
8. Health									N	-	N	5.259
9. Citizenship										-	N	5.196
10. Mathematics											+	9.250
11. Science												6.893

Key:

- \- indicates curriculum area in row ranked significantly lower than curriculum area in column;
- N indicates rankings of curriculum area in row and curriculum area in column were not significantly different;
- \+ indicates curriculum area in row ranked significantly higher than curriculum area in column;

Type I error rate (α) = 0.05.

Friedman's $\chi^2_r = 277.49$; $_{.95}\chi^2_{10} = 18.31$

Table 8. Results of Friedman's Test and Nonparametric Analog of Tukey's Studentized Range Tests Applied to Average Ranks Assigned Eleven Curriculum Areas by 56 Respondents. Based on Presentation of Activity Type Statements in Likert Scale Format.

Curriculum Area	1	2	3	4	5	6	7	8	9	10	11	Average of 56 Ranks
1. Music		N	-	N	N	N	N	N	N	-	N	4.420
2. Art			N	N	N	N	N	N	N	N	N	5.366
3. Reading				N	N	N	N	N	N	N	N	6.830
4. Productive Language					N	N	N	N	N	N	N	6.571
5. Receptive Language						N	N	N	N	N	N	5.625
6. Safety							N	N	N	N	N	6.505
7. Physical Education								N	N	N	N	5.893
8. Health									N	N	N	6.152
9. Citizenship										N	N	5.938
10. Mathematics											N	6.830
11. Science												5.830

Key:

- \- indicates curriculum area in row ranked significantly lower than curriculum area in column;
- N indicates rankings of curriculum area in row and curriculum area in column were not significantly different;
- \+ indicates curriculum area in row ranked significantly higher than curriculum area in column;

Type I error rate (α) = 0.05.

Friedman's $\chi^2_r = 25.92$; $_{.95}\chi^2_{10} = 18.31$

assign equal average ranks to all eleven stimulus statements. This null hypothesis was rejected well beyond the .001 level. The body of Table 5 shows the results of nonparametric analogs to Tukey's Studentized Range tests, using an experimentwise Type I error rate of .05. We conducted simultaneous tests on all 55 pairwise contrasts among the eleven stimulus statements. Significant differences were found for 34 of the 55 pairs tested. Respondents' preference for music was significantly lower than their preferences for all other curriculum areas except art, safety, and physical education. Art was ranked significantly lower than reading, productive language, receptive language, citizenship, mathematics, and science. There were no other significant differences between the ranking of art and other curriculum areas. Reading was ranked significantly higher than all other curriculum areas except productive language, receptive language and mathematics. In terms of average ranks, it was the most highly preferred of the curriculum areas. Productive language was preferred significantly more than all other curriculum areas except reading, receptive language, and mathematics. There were no significant differences between preferences for receptive language and preferences for reading, productive language, citizenship, mathematics, and science. However, receptive language was preferred significantly more than all remaining curriculum areas. For activity-type statements presented in the paired-choice format, preferences for the remaining curriculum areas differed among themselves in both directions. That is, safety, physical education, health, citizenship, and science were all in the middle range of preferences, and generally not reliably distinguishable from each other in terms of preference ordering.

Another statistic that indicates subjects' abilities to distinguish among preferences for various curriculum areas (or to agree with each other on a single preference ordering) is Kendall's coefficient of concordance. For

activity-type statements presented in the paired-choice format, Kendall's statistic was 0.463. This value is substantially greater than zero (which would indicate no agreement at all) and is in fact statistically significant at the .05 level; yet it is far below a value of one, which would indicate perfect agreement among respondents.

Table 6 contains results that parallel those of Table 5, for activity statements presented in the paired-choice format. Again, the null hypothesis of equal preference ranks was rejected using Friedman's test. The test statistic had a value of 267.48, which was significant well beyond the .001 level. Thirty-two of the 55 pairwise contrasts between stimulus statements tested for significance with Tukey's Studentized Range Procedure were found to be significant at the .05 level. Comparison of Tables 5 and 6 suggests great similarity in the patterns of significant differences between preferences for curriculum areas using activity-type statements and the more explicit activity statements as stimuli. Although there were some differences in the respondents' preference orderings using the two types of stimulus materials, the five top ranked curriculum areas were identical for both. Citing our earlier example of a hypothetical policy response to these data, in this case the school system would take identical action on the basis of survey results for both types of stimulus materials; additional resources would be placed in the same five curriculum areas.

For activity statements presented in the paired-choice format, Kendall's coefficient of concordance equaled 0.478, slightly higher than the value reported for activity-type statements. This indicates slightly greater agreement among respondents on a particular preference ordering of curriculum areas. However, the difference is probably not significant, either statistically or educationally.

Results of Friedman's test and nonparametric analogs of Tukey's Studentized Range tests for activity statements presented with accompanying photographs are very similar to results discussed previously for activity statements alone and for activity-type statements. Friedman's chi-square for activity statements with accompanying photographs in the paired-choice format, was 277.49, highly significant beyond the .001 level. With a few exceptions for the mid-ranked stimulus statements, patterns of significant contrast are virtually identical for activity statements presented in the paired-choice format, with or without accompanying photographs. Thirty of 55 possible contrasts were statistically significant at the .05 level for activity statements with photographs. The stimulus statements given the top five ranks were, in fact, identical for activity statements with or without photographs and for activity-type statements in the paired-choice format. Thus, the detailed description of results presented in Table 5 for activity-type statements applies in large part to the detailed results in Table 7 for activity statements with accompanying photographs.

Kendall's coefficient of concordance was 0.496 for activity statements with photographs presented in the paired-choice format. Thus, respondents presented with more-concrete stimulus materials (longer statements and photographs of actual activities in classrooms) did show somewhat greater ability to agree on a particular preference ordering for the eleven curriculum areas. Or, put another way, respondents' abilities to discriminate between their preferences for the various curriculum areas were somewhat greater. These results are consistent with our research hypothesis that increasing the concreteness of stimulus materials would reduce error variance in respondents' stated preferences and thus lead to more reliable results. However, this trend is not strong and may not be a statistically significant finding.

Results of Friedman's test and nonparametric analogs of Tukey's Studentized Range test applied to contrasts between preferences for pairs of activity-type statements presented in a Likert scale format are shown in Table 8. These results differ substantially from those shown in Tables 5 through 7. Friedman's chi-square statistic equaled 25.92. Although this value was statistically significant at the .05 level, it was about a tenth the size of corresponding statistics for materials presented in the paired-choice format. Thus, discrimination between preferences for different curriculum areas was not nearly as great when stimulus materials were presented in the Likert scale format. Results of pairwise contrasts provide further evidence in the support of this contention. Only two of 55 possible contrasts proved to be statistically significant at the .05 level. We can safely reject the null hypotheses that reading and music are equally preferred and that music and mathematics are equally preferred in the population represented by the 56 sampled respondents. However, no other reliable differences could be supported on the bases of these data. Kendall's coefficient of concordance for activity type statements presented in the Likert format was only 0.046, a value that is very close to zero. The inferiority of the Likert format for discriminating among respondents' preferences for various curriculum areas is once again supported.

Our discussion of data presented in Tables 5 through 7 has already suggested a high degree of similarity among preference orderings of curriculum areas for materials presented in the paired-choice format, regardless of the particular stimulus materials used. This earlier observation is further supported by the data presented in Table 9. Here we see that Spearman rank-order correlations are quite likely as high as the reliabilities of the individual presentation procedures will allow. As one would expect, the methods that employ stimulus materials that are more similar in level of

Table 9. Spearman Rank-order Correlations Between Preference Orderings of Eleven Curriculum Areas, Resulting from Various Paired-Choice Stimulus Formats.*

	Activity Statements	Activity Statements With Photographs
Activity-Type Statements	.964	.936
Activity Statements		.973

* Ranks averaged across 56 respondents per format

concreteness show slightly higher relationships than do the methods that use materials that are least similar. The materials that differ most are activity-type statements and activity statements with accompanying photographs. These data lend some credence to our hypothesized ordering of stimulus materials along a concreteness dimension.

In summary then, our research hypotheses have generally been supported. First, respondents' preferences for stimulus materials presented in the paired-choice formats and stimulus materials presented in the Likert format were substantially different. Second, all paired-choice formats were substantially more effective than the Likert scale format in allowing respondents to discriminate among their preferences for various curriculum areas. Third, use of more concrete and explicit stimulus materials in the paired-choice format appeared to enhance respondents' abilities to discriminate among their preferences for various curriculum areas. And finally, variation of stimulus materials in the paired-choice format did not result in substantial differences among respondents' preference orderings of curriculum areas.

IV. Implications for Practice in School Systems' Surveys

Although the research questions examined in this study might appear to be solely of theoretical interest, the findings of this research are of practical value to school systems. Three issues are most apparent: First, can school systems feasibly conduct the kind of household interview survey employed here to determine parents' or citizens' preferences for various elements of their curricula? Second, should school systems continue to use stimulus methods that incorporate a Likert scale for response, as was found to be typical practice in the literature reviewed, or should they use materials that incorporate a paired-choice format? Third, is there any advantage to using stimulus

materials that describe curriculum elements in concrete terms, or materials that incorporate visual as well as verbal elements? We shall examine these questions in order.

A. Feasibility

In order to determine whether household interview surveys of the type used in this study would be feasible for most school systems, we must consider such factors as cost, availability of personnel to conduct such surveys, and the likely effectiveness of such surveys as vehicles for securing the views of a representative sample of parents or citizens.

In general, interview surveys cost more than surveys conducted by mail. Interviewers must be recruited, trained, and usually paid for their time. Whether these costs are prohibitive depends on several factors, such as the local labor market and the availability of personnel with survey expertise within the school system's staff. If the school system is near a college or university, students can often be recruited and used to conduct interviews at relatively low cost. Payment of three to five dollars per interview, with the interviewers providing their own transportation, is not at all unreasonable. Another option was reported in a study conducted by Heath (1972). In that study, a professional research company was given a contract to plan, design, and manage a survey of citizens in a California school system, but interviewers were recruited by the school system through the local parent-teachers association, and worked on a volunteer basis at no cost. Interviewers were trained and monitored by the professional research company.

If a school system had to hire professionals to plan and design a survey, and then hire interviewers to conduct the survey, costs would depend largely on the size of the sample of respondents. If 200 parents or citizens were interviewed, and they were selected at random from a population at least ten

times as large, the percentage of the population that favored a given curriculum element, or desired a given school system practice, could be estimated within seven percent with 95 percent confidence. That is, the sample percentage and the true population percentage might differ by as much as seven points, but would not likely differ by more than that amount. Using professionals for all survey planning and design, and paying interviewers at the rate of four dollars per interview, a school system might spend about $2000 to $3000 to conduct such a survey. Whether such an expenditure was feasible would probably depend on the size of the school system.

The survey conducted for this study achieved a response rate of 81.5 percent. That is the percentage of parents who were contacted, who actually provided useful survey data. This figure compares very favorably with the 34.3 percent to 41 percent response rates reported in the ERIC Documents File referenced above, for school system surveys conducted by mail.

Low survey response rates threaten the validity of survey results in two ways. First, they increase random error by effectively reducing sample size. With a smaller sample size, the difference between an estimate for a population of parents or citizens, and the actual value for the population, may be much greater than that resulting from a larger sample. Second, and of greater concern, low survey response rates bias survey results systematically. Those who fail to respond to mail surveys tend to be different from those who do respond, in important and predictable ways: They are from lower socio-economic classes, more likely to be members of minority ethnic groups, have fewer years of schooling, and in the case of school system surveys, have less interest in the schools. Omission of such parents or citizens would definitely bias survey results to a substantial degree. A school system that achieves a

30 to 50 percent response rate in its survey cannot claim that it has assessed the desires or preferences of a representative sample of citizens or parents. In short, interview surveys may be required to achieve adequate survey results, and may be the only feasible alternative for school systems intent on doing an effective job of determining their constituencies' preferences.

B. Choice of Stimulus Method

It is evident from the results of this study that the paired-choice stimulus method provides results that are far more useful for decision-making purposes than does the Likert scale format. As described in Section III, respondents presented with stimulus materials in the paired-choice format were in general agreement on their relative preferences for various curriculum areas, regardless of the type of stimulus materials used. In addition, their ranking of curriculum preferences differed substantially from the preference ranking produced by respondents to the Likert scale format. Detailed analyses revealed that respondents to materials presented in the paired-choice format provided many statistically significant discriminations among their preferences for various curriculum areas. In contrast, respondents to materials presented in the Likert scale format generally wanted the school system to do everything at once; they provided few statistically reliable preference discriminations.

One potential problem with the paired-choice format is the length of time required to complete an interview using this method. In this study, interviews using activity-type statements in the paired-choice format required just under an hour to complete. When the same stimulus materials were used in the Likert scale format, interviews averaged only fifteen minutes. When the longer activity statements were used in the paired-choice format, interviews required more than an hour.

The paired-choice format requires substantially more time than the Likert scale format because each stimulus statement must be paired with every other, and the number of pairs increases very rapidly as the number of stimulus statements is increased. In this study, use of eleven stimulus statements resulted in 55 pairs. Twelve stimulus statements would have resulted in 66 pairs; thirteen stimulus statements in 78; fourteen in 91, etc. One way of avoiding this problem would be to increase the sample size of respondents, and not present every respondent with every pair of stimulus statements. For example, if pairs of stimulus statements were assigned to respondents in accordance with a balanced incomplete block experimental design, and if the size of the respondent sample were to be doubled, interview time could be cut by almost 50 percent since each respondent would be presented with only half the possible pairs of stimulus statements.

C. Concreteness of Stimuli

The results of this study show that increasing the concreteness of stimulus materials enables respondents to discriminate among their preferences for various curriculum areas to a slightly greater degree. However, the increase in discrimination is not substantial, and preference rankings of the curriculum areas barely differed among the types of stimulus materials presented in the paired-choice format. Thus a school system that used shorter, more-general, activity-type statements would likely achieve the same preference ranking as would a school system that used more expensive and time-consuming stimulus materials, composed of detailed activity statements with accompanying photographs.

In summary then, a school system planning to conduct a survey of parents or citizens would be best advised to adopt the following procedures:

1) Conduct a household interview survey rather than a mail survey; 2) use short, totally verbal descriptions of types of activities that are representative of the school system's curriculum; 3) present stimulus materials in a paired-choice format, rather than in a Likert scale format. This combination of design features, together with the adoption of other sound survey procedures -- such as random sampling of respondents, adequate training of interviewers, attention to the quality of data through careful editing, etc. -- should provide school systems with the highest quality survey results at least cost.

REFERENCES

Ahmann, S. A report on national assessment in seven learning areas. Today's Education, 1975, 64, 63-64.

Arizona State University, Tempe. Report of the state educational needs assessment project of Arizona. Tempe, Arizona, 1969. (ERIC Document Reproduction Service No. ED 080 619)

Beverly Hills Unified School District. Educational needs assessment: A community survey. Beverly Hills, California, 1972. (ERIC Document Reproduction Service No. ED 072 073)

Blackwell, P. J. Parental viewpoints of goals in secondary education: The methodological procedures. Washington, D.C.: National Institute of Education, 1973. (ERIC Document Reproduction Service No. ED 088 933)

Bucks County Public Schools, Doylestown, Penn. General needs assessment: Quality educational program study booklet B. Harrisburg: Penn. State Dept. of Education, 1971. (ERIC Document Reproduction Service No. ED 063 357)

Crim, R. D. et al. Needs assessment for staff development. Luconia, New Hampshire: New Hampshire Supervisory Union 30, 1975. (ERIC Document Reproduction Service No. ED 106 271)

Eisner, E. Emerging models for educational evaluation. School Review, 1972, 80, 573-590.

Eugene School District 4J. A report on parent opinion surveys. Eugene, Oregon: Oregon Division of Research, Development and Evaluation, 1975. (ERIC Document Reproduction Service No. ED 128 964)

Faunce, R. W. et al. An approach to setting educational goal priorities in the Minneapolis Public Schools. Minneapolis, Minnesota: Minneapolis Public Schools, 1974. (ERIC Document Reproduction Service No. ED 117 163)

Gilbert, K. & Ellis, E. N. A survey of parents' opinions regarding school policies and practices at John Oliver Secondary School. Vancouver: Vancouver Board of School Trustees, 1971. (ERIC Document Reproduction Service No. ED 077 993)

Greensboro Public Schools. Comprehensive plan for instruction grades k-12. Greensboro, North Carolina: 1977.

Heath, R. W. Preferences for educational goals in San Bruno. Berkeley, California: Nomos Institute, 1972. Unpublished manuscript.

Horner, P. J. & Martin, H. Needs assessment survey: The school image. Shiprock, New Mexico: Shiprock independent school district 22, 1977. (ERIC Document Reproduction Service No. ED 149 942)

McGuire, T. C. & Petrosko, J. M. Determining educational goal priorities in a secondary school. Los Angeles: University of California at Los Angeles. Center for the Study of Evaluation, 1974. (ERIC Document Reproduction Service No. ED 098 196)

McTeer, J. H. et al. A study of the differences of views of students, parents, teachers, and administrators on selecting teaching objectives. West Georgia College, 1976. (ERIC Document Reproduction Service No. ED 130 957)

Moser, C. A. & Kalton, G. Survey methods in social investigation. New York: Basic Books, Inc., 1972.

Moses, L. E. & Oakford, R. V. Tables of random permutations. Stanford: Stanford University Press, 1963.

National Study of School Evaluation. Parent opinion inventory. Arlington, Virginia: 1976. (ERIC Document Reproduction Service No. ED 155 183)

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. Statistical package for the social sciences, 2nd ed. New York: McGraw-Hill, 1975.

Preising, P. P. A survey of the educational needs of Santa Clara County. San Jose, California: Supplementary Education Center, 1967. (ERIC Document Reproduction Service No. ED 018 509)

Rankin, R. J. et al. Educational needs as perceived by seven constituencies in education. Paper presented at the Annual Meeting of the AERA, San Francisco, 1976. (ERIC Document Reproduction Service No. ED 123 769)

Rhode Island State Department of Education. Needs assessment compendium of abstracts. Providence: 1978. (ERIC Document Reproduction Service No. ED 153 351)

Rookey, T. J. Needs assessment model: East Stroudsburg. East Stroudsburg, Pennsylvania: East Stroudsburg State College, 1976. (ERIC Document Reproduction Service No. ED 133 828)

SAS Institute. SAS user's guide, 1979 ed. Raleigh, N.C.: 1979.

Saskatoon Separate Board of Education. Citizens' school inquiry. Saskatoon, Saskatcheuan, 1973. (ERIC Document Reproduction Service No. ED 078 528)

Savage, J. & Jones, J. Parents concerns with education. Elementary School Journal, 1973, 73, 125-129.

Sawyers, B. A. Assessment and evaluation of public's perception of vocational education in Indiana. Indianapolis, Indiana: Indiana State Board of Voc. & Tech. Education, 1976. (ERIC Document Reproduction Service No. ED 136 015)

Tacoma School District Ten. Education for the 70's: A survey of community opinion about the Tacoma Public Schools. Tacoma, Washington, 1971. (ERIC Document Reproduction Service No. ED 053 191)